Journal of Life Sciences & Biomedicine, vol. 1(74), No 1, p. 37-46 (2019)

Study of genomic variation in bread wheat collection based on next generation sequencing data

Z.İ. Akparov, M.A. Abbasov*

*Genetic Resources Institute, Azerbaijan National Academy of Sciences, 155 Azadlig ave., Baku AZ 1106, Azerbaijan; *For correspondence:mehraj_genetic@yahoo.com*
Accepted for publication: 18 September 2019

Genotyping by Sequencing (GBS) is a Next Generation Sequencing (NGS) technique widely applied in plant breeding that uses restriction enzymes to reduce the complexity of the genome. In the current study the genomic diversity of 87 local and introduced bread wheat genotypes was evaluated using GBS technology. A total of 411 single-nucleotide polymorphisms (SNPs) were obtained for three genomes. The SNP range within each genome was 15–29, 10–36 and 3–17 for A, B and D genome, respectively. The highest number of SNP markers was recorded on the B (48.8%) and the lowest on the D genome (14%). In total, 70.2% of SNPs were transitions (Ts) and 29.8% transversions (Tv). The largest Delta K value was recorded at K = 2, indicating the existence of 2 groups in the collection. The I group contained 68.5% of the introduced accessions, whereas 82% of local genotypes fell into the II group. The average ancestral contribution of the genotypes in I and II groups were 86.4% and 83.6%, respectively. The results of cluster and PCoA analyses were consistent with the STRUCTURE, indicating a sharp differentiation between local and introduced germplasm. Other factors determining the grouping of samples were traits of botanical varieties and genealogy. The SNP markers, revealed in the current study will be used as a genetic source for genotyping and marker-trait association analyzes. The data can be successfully applied in the development and implementation of new strategies for subsequent genetic analysis and breeding.

Keywords: *Bread wheat, genome, genotyping by sequencing, SNP, transition, transversion*

INTRODUCTION

Bread wheat (*Triticum aestivum* L.) (AABBDD) is one of the most important cereal crops ensuring food security at the global level (Tadesse, 2016). It occupies more than 90% of the total area cultivated with wheat and mainly used for preparing bread and biscuits. The production of wheat has dramatically increased and continues to increase to meet the steadily growing needs of the world population. To make this increase sustainable new genetic sources must be identified and involved into the breeding process. Wheat genetic resources are the main sources in wheat improvement programs. To date, more than 900,000 wheat accessions, including wild relatives, landraces and breeding lines are conserved in different genebanks at the global level. Management of these genetic resources is a big challenge and requires the use of modern strategies, such as genomic tools and techniques (Tadesse, 2016). The whole genome of bread wheat (Chinese Spring [CS42]) was sequenced in 2012 by Brenchley and co-workers with the Roche 454 Next Generation Sequencer (NGS) and 94,000-96,000 genes were identified in its 17 GB genome. The gradual decline in the financial cost of the sequencing technologies has facilitated the use of the NGS and prompted the creation of new genotyping methods, such as genotyping by sequencing (GBS). GBS is the most successful genotyping method applied to plant breeding (Poland and Rife, 2012; Singh et al., 2019). The method is used in a wide range of breeding programs, including genome selection, Genome Wide Association Mapping, and new marker detection (Poland and Rife, 2012; Narum et al., 2013). It is a high-throughput multiplex genotyping technique that uses restriction enzymes

to reduce the complexity of the genome. GBS has a number of advantages to investigate polyploid species such as bread wheat. The wheat genome is composed of 80% repeated sequences, which can be avoided using methylation-sensitive restriction enzymes during GBS. Due to barcoding, GBS allows the genotyping of hundreds of samples at the same time, depending on the platform (Elshire et al., 2011). Genomic data obtained during genotyping by sequencing are equivalent to the information provided by thousands of molecular markers.

Azerbaijan is considered to be one of the major origin countries and has a large biodiversity of wheat and its wild relatives. This biodiversity has been collected for many years, characterized and evaluated by various methods and involved into breeding processes. As a result, hundreds of landraces and new varieties have been created, a large portion of which along with collected and introduced accessions have been conserved at the National Genebank in the Genetic Resources Institute. Despite various molecular techniques were used to study genetic diversity in this germplasm, the application of NGS technologies, which is considered the most modern genomic approach worldwide is very relevant (Abbasov et al., 2018). Thus the main objective of the study is to characterize the allelic and genetic diversity of local and introduced bread wheat genotypes from National Genebank using NGS-based GBS technology.

MATERIALS AND METHODS

Eighty-seven local and introduced bread wheat accessions conserved at the National Genebank of the Genetic Resources Institute of ANAS were used in the study. The list and characteristics of the studied accessions are given in Table 1.

Table 1. The list of used bread wheat genotypes

№	Botanical variety	№	Botanical variety	№	Botanical variety
6847	v. *erythrospermum*	6983	v. *lutescens*	10425	
6920	v. *graecum*	6984	v. *lutescens*	10430	
6926	v. *graecum*	6985	v. *lutescens*	20125	
6927	v. *graecum*	6987	v. *lutescens*	20783	
6936	v. *milturum*	6989	v. *lutescens*	20785	
6937	v. *milturum*	6992	v. *lutescens*	21134	
6938	v. *milturum*	7008	v. *albidum*	21139	
6939	v. *milturum*	7010	v. *albidum*	21338	
6941	v. *milturum*	7012	v. *albidum*	21339	
6942	v. *milturum*	7014	v. *albidum*	Akinchi 84	v. *erythrospermum*
6944	v. *erythrospermum*	7016	-	Azametli 95	v. *graecum*
6945	v. *erythrospermum*	7017	-	Azeri	v. *lutescens*
6948	v. *erythrospermum*	7019	v. *barbarossa*	Chinese spring	
6949	v. *erythrospermum*	7020	v. *barbarossa*	Giymetli2/17	v. *velutinum*
6950	v. *erythrospermum*	7021	v. *barbarossa*	Graekum75/50	v. *graecum*
6959	v. *alborubrum*	7023	v. *barbarossa*	Guneshli	v. *erythrospermum*
6960	v. *alborubrum*	7027	v. *hostianum*	Gurgene 1	v. *erythrospermum*
6961	v. *alborubrum*	7028	v. *hostianum*	Jagger	
6962	v. *alborubrum*	7029	v. *hostianum*	Mirbashir 128	v. *erythrospermum*
6963	v. *alborubrum*	7032	v. *meridionale*	Morocco	
6964	v. *ferrugineum*	7033	v. *meridionale*	Girmizigul	v. *erythrospermum*
6965	v. *ferrugineum*	7034	v. *leucospermum*	Gobustan	v. *graecum*
6966	v. *ferrugineum*	7036	v. *velutinum*	Ruzi 84	v. *graecum*
6968	v. *ferrugineum*	10377		Shafaq	v. *lutescens*
6969	v. *ferrugineum*	10378		Sheki1	v. *lutescens*
6970	v. *ferrugineum*	10380		Taraqqi	v. *lutescens*
6971	v. *ferrugineum*	10381		Turkey	
6972	v. *ferrugineum*	10383		Yegane	v. *ferrugineum*
6973	v. *ferrugineum*	10384		Zerdabi	

The DNA from young leaves of each accession was extracted using the CTAB procedure (Doyle and Doyle, 1987). The genomic DNA was co-digested with the restriction enzymes *Pst*I (CTGCAG) and *Msp*I (CCGG) (New England BioLabs, Inc., Ipswich, MA, USA) after which barcoded adapters were ligated using T4 ligase. Samples were pooled by plate into libraries, purified using the QIAquick PCR Purification Kit (Qiagen, Inc., Valencia, CA, USA) and polymerase chain reaction amplified. The PCR products were again cleaned up using the same kit, and size-selected in an E-gel system (Life Technologies, Inc.). Each 95-plex library was sequenced on Illumina HiSeq 2500. Sequence results were analyzed using the UNEAK GBS pipeline, TASSEL 3.0. The obtained single nucleotide polymorphism (SNP) markers were filtered and SNPs with more than 40% missing data and with a minor allele frequency (MAF) less than 10% were excluded from the analysis. Nei genetic distances among individuals and botanical varieties were calculated using PoweMarker V3.25 software (Liu and Muse, 2005). PCoA and cluster analyses based on Nei genetic distances (Nei, 1972) among individuals and botanical varieties were performed using DARwin 6.0 (Perrier and Jacquemoud-Collet, 2006). Population structure analysis was done using ADMIXTURE software.

RESULTS AND DISCUSSION

A total of 411 single-nucleotide polymorphisms (SNPs) were obtained for three genomes (ABD) based on GBS analysis of 87 local and introduced wheat varieties and accessions. The distribution pattern of SNPs over the three genomes differed, with the highest number of markers on the B (48.8%) and the lowest marker number on the D genome (14%) (Figure 1). This could be due to the relatively recent introgression of the D genome.

The distribution pattern of SNP markers across the A, B and D genomes is consistent with previous studies (Akhunov et al., 2010; Marcussen et al., 2014; Shavrukov et al., 2014; Edae et al., 2015). Pour et al. (2017) genotyped 369 Iranian hexaploid wheat accessions using 16506 GBS-derived SNPs and found that most of the SNPs were located on the B genome, while the D genome had the least number of markers. The uneven distribution of genetic variation over the *T. aestivum* genome is associated with reduced genetic recombination due to the self-pollination and genetic mechanisms that prevent the pairing of homoelogical chromosomes during meiosis.

The A/D or B/D SNP ratio in the current study was lower than several other studies (Allen et al., 2011; Cavanagh et al., 2013), indicating that the bread wheat accessions of Azerbaijan have a relatively higher SNP variation on the D genome. The presence of D genome largely determines the high baking properties of bread wheat. The high diversity found on the D genome ensures that the collection can be used as a source of new, desirable alleles for traits of agronomic importance, including high baking quality (Jia et al., 2013).



Figure 1. Number of SNP markers mapped on the hexaploid wheat genome.

The number of SNPs on each chromosome varied widely (3-36), with a maximum number on chromosome 5B and minimum on chromosome 5D. The SNP range within each genome was 15–29, 10–36 and 3–17 for A, B and D genome, respectively. Within each genome the least number of SNPs was observed on the homoeologous group 4. Chromosomes of homoeologous group 4 are relatively conservative, with some significant genes located there; for example, one recessive gene responsible for male sterility is located on the short arm of the chromosome 4B (Barlow and Driscoll, 1981). Therefore, any large variation or mutation that occurs on these chromosomes can cause the death of plants. In addition, homoeologous group 4 has a lower number of genes than the remaining homoeologous groups. Since recombination occurs primarily in genes, the low number of genes on this group may also lead to a low frequency of crossover and subsequently to the decrease of polymorphism and variation (Dvorak and McGuire, 1981).

In total, 70.2% of SNPs were transitions (Ts) and 29.8% transversions (Tv). The majority of SNPs were A→G and C→T transitions, with only 3 transitions in the reverse direction detected on the A genome. The highest number of transition-type SNPs was identified on the B (134 SNPs), while the lowest was on the D genome (42 SNPs). Transversion-type substitutions were in 5 directions (C/G, A/C, G/T, A/T and T/A); the only T/A transversion was detected on the B genome. A Ts/Tv ratio over the three genomes of hexaploid wheat was 2.36.

The most common mechanism of transition is the mutation of methylcytosine to uracil and its subsequent replacement with thymine. Genome-wide methylation is considered to be the direct result of polyploidy (Charmet, 2011). The high number of transitions observed on A and B genomes may be due to the two rounds of polyploidy and methylation events during the evolution of the hexaploid wheat (Buckler and Holtsford, 1996).

After filtration 313 SNP markers were used to evaluate the genetic diversity and population structure of bread wheat collection. Genetic diversity index (GDI) and the polymorphism information content (PIC) for 87 genotypes varied between 0.245-0.50 and 0.215-0.375; the average for the collection was 0.422 and 0.331, respectively. For the vast majority of SNP data, GDI is characterized by high values, and for 42% of SNPs maximum value was recorded. A similar tendency was also observed for the PIC value, indicating the presence of rich genome diversity in the collection, the majority (78.2%) of which comprised of Azerbaijani varieties and accessions.

Dvorak et al. (2006) suggested that compared to other places, *T. aestivum* in East Asia represents more original genepool than anywhere else. The spread of this genepool to eastern Turkey through the South Caucasus or south-west coast of the Caspian Sea, caused to the sympatry of the specie with wild emmer, the gene flow from wild emmer to *T. aestivum* and lead to the increase of the genetic variation and to changes in the geographic form of genetic diversity in *T. aestivum*.

The highest polymorphism among the bread wheat botanical varieties was recorded in var. *ferrugineum*, while the least polymorphism was in var. *hostianum* (Table 2). In general, the PIC value depended on the number of samples per botanical varieties.

Table 2. PIC values in bread wheat botanical varieites

Botanical varieties	Sample size	PIC
var. *albidum*	4	0.175
var. *alborubrum*	5	0.203
var. *barbarossa*	4	0.190
var. *erythrospermum*	12	0.276
var. *ferrugineum*	10	0.283
var. *graecum*	7	0.217
var. *hostianum*	3	0.075
var. *lutescens*	10	0.243
var. *meridionale*	2	0.163
var. *milturum*	6	0.245
var. *velutinum*	2	0.111

The delta K value was used to identify subpopulations in a bread wheat collection. The largest Delta K was recorded at K = 2, indicating the existence of 2 groups in the collection (Figure 2). The I group called "Introduction Group" consisted of 25 accessions (red stripe), 13 of which were samples from different countries and 12 were local genotypes. In total, 68.5% of the introduced accessions and only 2 out of 20 varieties (Chinese spring and Morocco) were included in this group. The average ancestral contribution of the genotypes was 86.4%. Of the 25 accessions, five, including Morocco variety had a high proportion of admixture.

The second group (Local Group) contained the remaining 62 accessions (green stripe), including 6 introduced genotypes and 18 varieties. The average ancestral contribution of the samples was 83.6%. Genetic diversity within the Local group (GDI=0.385; PIC=0.305) was higher than the Introduced group (GDI=0.336; PIC=0.268).

The genetic relationship among bread wheat accessions was further studied based on the cluster (Figure 3) and PCoA analyses (Figure 4). Among the samples, the Nei Genetic Distance Index (GD) index varied between 0.0-1.0 and averaged 0.57. Based on the SNP data, 100% similarity was observed between groups of genotypes, while the most genetically remote genotypes were 6926 and 20125. The average genetic distance between varieties and genebank accessions was 0.047. This indicates that there is no sharp differentiation between the two groups. The highest similarity among the varieties was recorded between the Azamatli 95 and Gobustan (GD=0.27), and the least between Mirbashir 128 and the Morocco variety (GD=0.776).

In a dendrogram developed by cluster analysis, bread wheat accessions were grouped into 3 clusters. In general, the results of the cluster analysis were consistent with the STRUCTURE analysis. Thus, the first cluster consists of 27 accessions and resembles the "Introduction group". Similarly, 13 of the 19 introduced accessions, including the Chinese spring and Morocco varieties, were grouped in this cluster. Within the group, the genotypes from the same country (e.g., Iran and Afghanistan) were placed closer.

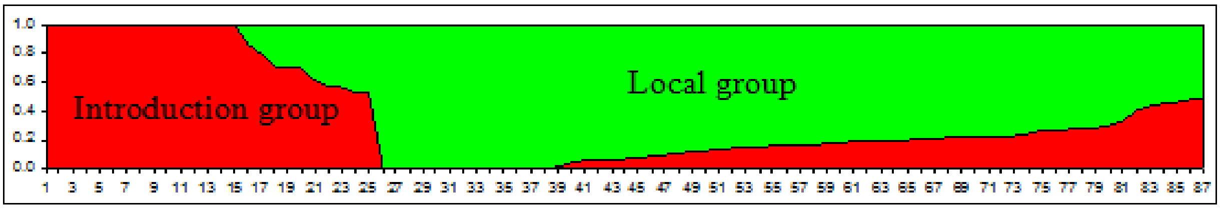


Figure 2. STRUCTURE analysis of bread wheat collection

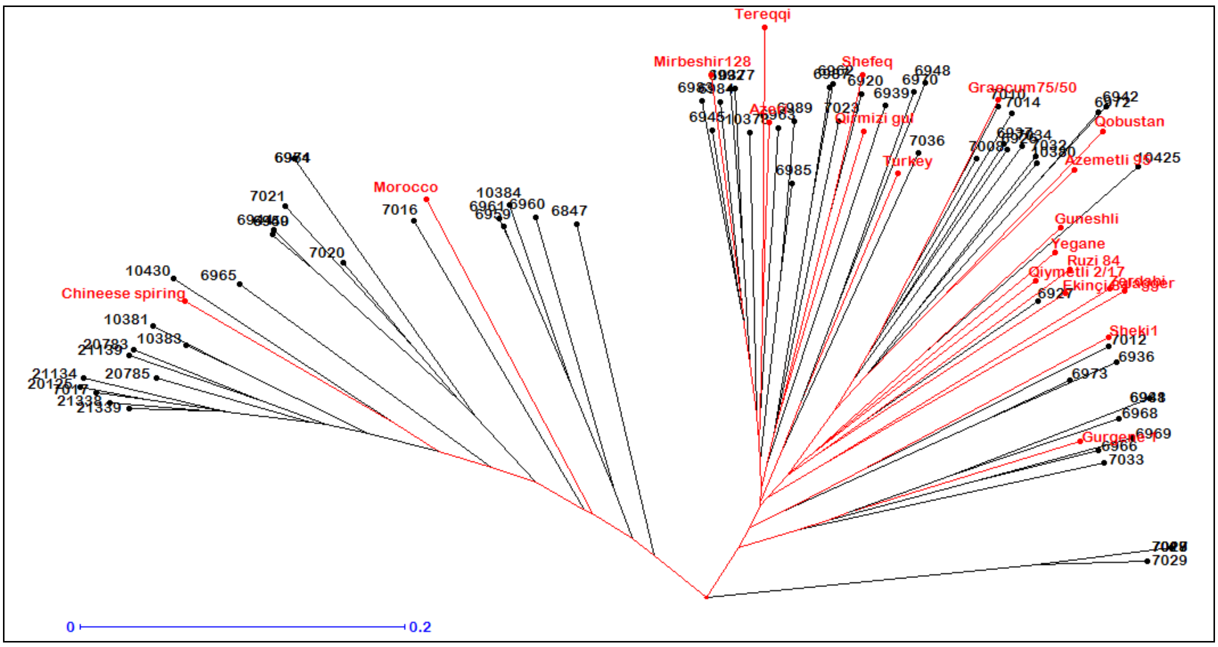


Figure 3. The dendrogram representing the genetic relationship among *T. aestivum* accessions based on GBS-SNP data. The varieties are given in red.

Cluster II consisted of 50 local and only 6 introduced genotypes, as was in "Local group" in Structure analysis. All local (16 varieties) and 2 foreign varieties (Turkey and Jagger) were also included into the cluster.

Finally, the III cluster contained only four genotypes. The genotypes included into the cluster had the highest admixture proportion according to STRUCTURE analysis. Of the four samples, three showed complete similarity.

Several irregularities can be underlined in the grouping pattern of genotypes in the dendrogram.

a) Differentiation of introduced and local accessions. The results of both cluster and Structure analyzes were able to differentiate between local and introduced accessions; with the exception of one accession, the introduced genotypes formed a homogenous subcluster in cluster I. Although 6 introduced accessions were included in another cluster, only two of them were genetically close to the local genotypes.

b) Joint grouping of genotypes of the same botanical varieties. There was a tendency of close grouping for most of the botanical varieties involved in the study. Thus, all 3 var. *hostianum* accessions were tightly grouped in cluster III and 100% similarity was observed between the 2 of them. In addition, 7 out of 10 var. *lutescens* genotypes, 5 out of 6 var. *graecum*, 3 out of 4 var. *alborubrum* and 3 out of 4 var. *albidum* genotypes were located in separate sub-clusters of clusters II and III. The joint grouping was also noted for var. *barbadosa*, var. *milturum* and other botanical varieties. Similar analyses have not been conducted since the botanical variety of introduced genotypes was unknown. The taxonomy on botanical variety had a significant impact on the overall view of the dendrogram.

c) Joint grouping of varieties. Twenty varieties, including 16 local ones were used in the current study, of which 18 were grouped in cluster II (group II) as a result of both analyzes. Although 1 of the 2 introduced varieties (Jagger) was grouped together with the local Zardabi variety, the genetic distance between them was quite high (0.42). The fact that the majority of varieties fell into the same cluster indicates the presence of common or shared alleles among them. The main priority in creating new varieties in local breeding programs was to obtain high-yielding and shorter (dwarf or semi-dwarf) varieties. Most of the studied accessions had short or medium (74-110 cm) height. It seems that the use of similar breeding programs and genetic resources has ultimately led to the creation of a genepool of local varieties with a common genetic background. However, no homogeneity was observed in the distribution of the species within the cluster, and they were placed in combination with genebank accessions.

d) Relationship between the grouping of varieties and genealogy. Some varieties with shared genealogy tended to group together. For example, the 3 species (Azeri, Taraggi and Mirbashir 128) were located in a separate subcluster of Cluster II, and the genetic distance index between the varieties varied from 0.3-0.36. Azeri and Taraggi varieties had common parent Panonia-45319, while Bezostaya-1 was used as a parent for the Azeri and Mirbashir 128 varieties. In addition, the Azamatli 95 and Gobustan varieties were obtained by individual selection from the ICARDA/CIMMYT genotypes.

There was no link between the grouping of the local samples and the geographical region. It can be assumed that the accessions were not grown for a long period in the regions from where they were collected, and the genepool here was formed as a result of the sowing and multiplication of seeds from certain sources by local people, as well as a result of seed exchange among neighboring regions. Therefore, the lack of the relationship between the genetic structure and the geographic region is expected. It should be reiterated that the samples introduced from the same country were found to be closer to each other genetically, indicating the use of similar and common genetic resources within those countries as well.

The distribution pattern of the bread wheat genotypes on the scatter plot was investigated by PCoA analysis, where the first three coordinates explained 25.7% of the SNP variation.

As shown in the figure, the grouping of the samples was in agreement with the previous analyzes. Varieties were located in the left and the introduced accessions in the lower right quadrant of the plot.

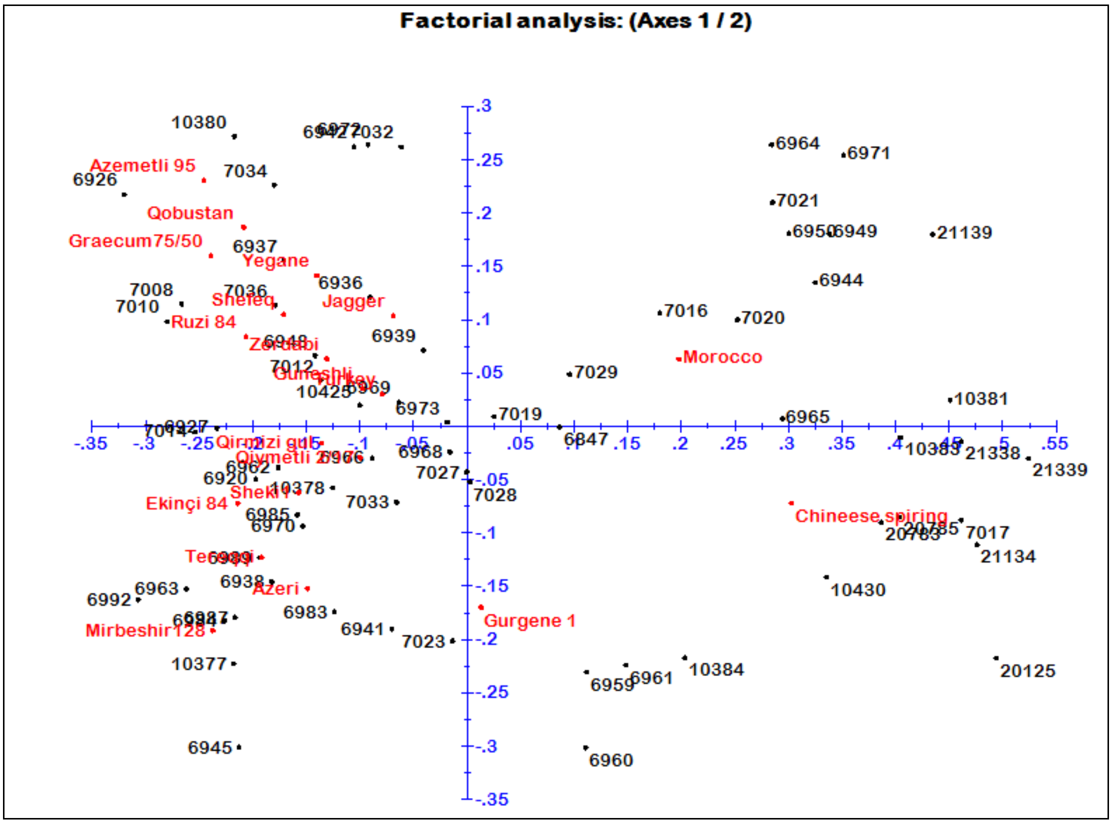


Figure 4. PCoA analysis of 87 bread wheat accessions.

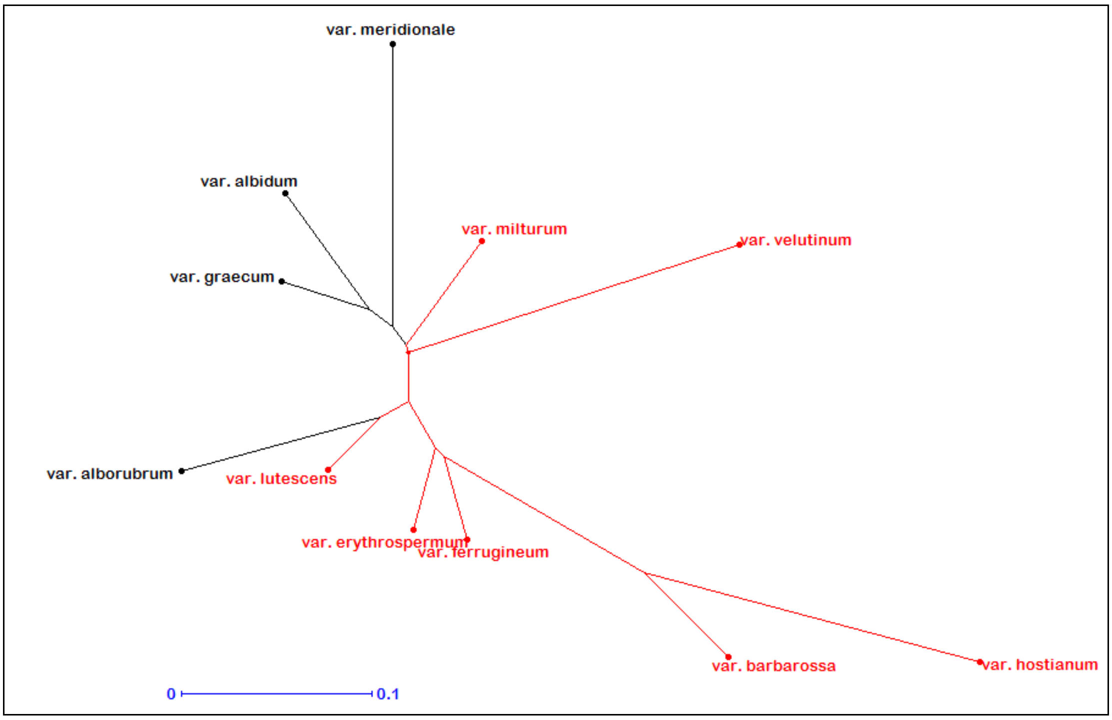


Figure 5. Dendrogram, based on SNP data, reflecting genetic relationship between bread wheat botanical varieties. The botanical varieties with red seeds is shown in red.

The Nei genetic distance index between botanical varieties varied from 0.11-0.68, with an average of 0.29. Var. *velutinum* and var. *hostianum* were the genetically most remote, while var. *ferrugineum* and var. *erythrospermum* have been found to be the closest.

Three clusters were found in the dendrogram representing the genetic relationship between botanical varieties. Var. *velutinum* formed an independent cluster (Figure 5). Unlike the others, the mentioned variety had hairy and white spikes without awns and red seeds.

Of the 6 botanical varieties in cluster II, 5 were red, whereas 3 of the 4 botanical varieties in cluster III had white seeds, suggesting that the seed color had a significant effect on the grouping of botanical varieties.

Summarizing the results of Cluster, STRUCTURE, and PCoA analyses, the use of seed material from the same sources in different geographical regions across the Republic for many years on one hand has resulted in their joint grouping with similar genetic background and, on the other hand, in differentiation from other introduced sources. Other factors determining the grouping of samples were traits of botanical varieties and genealogy. The lack of sharp differentiation between varieties and genebank accessions suggests that local populations are actively used in the breeding process.

So, based on SNP data, a large genomic variation was found among local and introduced bread wheat accessions. Wheat improvement programs are based on the use of molecular markers that significantly increase selective efficiency and allow the accurate transfer of genes and QTLs between different genetic sources. The SNP markers, revealed in the current study will be used as a genetic source and material for genotyping and marker-trait association analyzes. The data can be successfully applied in the development and implementation of new strategies for subsequent genetic analysis and breeding.

REFERENCES

Akhunov E.D., Akhunova A.R., Anderson O.D., Anderson J.A., Blake N., Clegg M.T., et al. (2010) Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genomics*, **11:** 702. doi: 10.1186/1471-2164-11-702.

Abbasov M., Akparov Z., Gross T., Babayeva S., Izzatullayeva V., Hajiyev E., Rustamov K., Gross P., Tekin M., Akar T., Chao S. (2018) Genetic relationship of diploid wheat (*Triticum* spp.) species assessed by SSR markers. *Genetic resources and crop evolution*, **65(5):**1441-1453.

Singh N., Wu S., Tiwari V.K., Sehgal S.K., Raupp J., Wilson D., Abbasov M., Gill B.S., Poland J. (2019) Genomic analysis confirms population structure and identifies inter-lineage hybrids in *Aegilops tauschii*. *Frontiers in plant science*, **10:** 9.

Alipour H., Bihamta M.R., Mohammadi V., Peyghambari S.A., Bai G., Zhang G. (2017) Genotyping-by-sequencing (GBS) revealed molecular genetic diversity of Iranian wheat landraces and cultivars. *Frontiers in plant science*, **8:** 1293.

Allen A.M., Barker G.L., Berry S.T., Coghill J.A., Gwilliam R., Kirby S., Robinson P., Brenchley R.C., D'Amore R., McKenzie N., Waite D. (2011) Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant biotechnology journal*, **9(9):** 1086-1099.

Barlow K.K., Driscoll C.J. (1981) Linkage studies involving two chromosomal male-sterility mutants in hexaploid wheat. *Genetics*, **98(4):** 791-799.

Brenchley R., Spannagl M., Pfeifer M., Barker G.L., D'Amore R., Allen A.M., McKenzie N., Kramer M., Kerhornou A., Bolser D., Kay S. (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, **491(7426):** 705.

Buckler E.S., Holtsford T.P. (1996) Zea ribosomal repeat evolution and substitution patterns. *Molecular Biology and Evolution*, **13(4):** 623-632.

Cavanagh C.R., Chao S., Wang S., Huang B.E., Stephen S., Kiani S., Forrest K., Saintenac C., Brown-Guedira G.L., Akhunova A., See D. (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proceedings of the national academy of sciences*, **110(20):** 8057-8062.

Charmet G. (2011) Wheat domestication: lessons for the future. *Comptes rendus biologies*, **334(3):** 212-220.

Doyle J.J., Doyle J.L. (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem Bull.*, **19:** 11–15.

Dvorak J., Akhunov E.D., Akhunov A.R., Deal K.R., Luo M.C. (2006) Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Molecular biology and evolution*, **23(7):** 1386-1396.

Dvořák J., McGuire P.E. (1981) Nonstructural chromosome differentiation among wheat cultivars, with special reference to differentiation of chromosomes in related species. *Genetics*, **97(2):** 391-414.

Edae E.A., Bowden R.L., Poland J. (2015) Application of population sequencing (POPSEQ) for ordering and imputing genotyping-by-sequencing markers in hexaploid wheat. *G3: Genes, Genomes, Genetics*, **5(12):** 2547-2553.

Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., Mitchell S.E. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, **6(5):** p.e19379.

Jia J., Zhao S., Kong X., Li Y., Zhao G., He W., Appels R., Pfeifer M., Tao Y., Zhang X., Jing R. (2013) *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*, **496 (7443):** 91.

Liu K., Muse S.V. (2005) PowerMarker: Integrated analysis enviornment for genetic marker data. *Bioinformatics*, **21:** 2128–2129.

Marcussen T., Sandve S.R., Heier L., Spannagl M., Pfeifer M. (2014) International Wheat Genome Sequencing Consortium Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, **345:** 1250092. doi: 10.1126/science.1250092.

Narum S.R., Buerkle C.A., Davey J.W., Miller M.R., Hohenlohe P.A. (2013) Genotyping-by-sequencing in ecological and conservation genomics. *Molecular ecology*, **22(11):** 2841-2847.

Perrier X., Jacquemoud-Collet J.P. (2006) DARwin software: Dissimilarity analysis and representation for windows. *Website http://darwin. cirad. fr/darwin*.

Poland J.A., Rife T.W. (2012) Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome*, **5:** 92–102. doi: 10.3835/plantgenome2012.05.0005.

Shavrukov Y., Suchecki R., Eliby S., Abugalieva A., Kenebayev S., Langridge P. (2014) Application of next-generation sequencing technology to study genetic diversity and identify unique SNP markers in bread wheat from Kazakhstan. *BMC plant biology*, **14(1):** 258.

Tadesse W., Amri A., Ogbonnaya F.C., Sanchez-Garcia M., Sohail Q., Baum M. (2016) Wheat. Academic Press. In: *Genetic and Genomic Resources for Grain Cereals Improvement*, 81-124.

Yeni Nəsil Sekvensləmə verilənləri əsasında yumşaq buğda kolleksiyasında genom variasiyasının tədqiqi

Z. İ.Əkpərov, M.Ə.Abbasov

AMEA Genetik Ehtiyatlar İnstitutu

Sekvens əsasında genotipləşdirmə (GBS) genom mürəkkəbliyini azaldılması üçün restriksiya fermentlərdən istifadə edən YNS genotipləşdirmə üsulu olub bitki seleksiyasında uğurla tətbiq olunur. Tədqiqat işində yerli və introduksiya olunmuş 87 yumşaq buğda genotiplərinin genom müxtəlifliyi YNS əsaslı GBS texnologiyasından istifadə edərək qiymətləndirilmişdir. Üç genom üçün ümumilikdə 411 tək nukleotid polimorfizm (TNP) markeri əldə edilmişdir. A, B və D genomu üzrə TNP miqdarı müvafiq olaraq 15–29, 10–36 və 3–17 intervalında dəyişmişdir. SNP markerlərinin ən çox sayı B (48,8%), ən az sayı isə D genomunda (14%) qeydə alınmış, SNP-lərin 70.2% -i tranzisiya (Ts), 29.8% -i isə transversiya (Tv) tipli olmuşdur. Delta K-nın ən böyük dəyəri K = 2 səviyyəsində qeydə alınmışdır ki, bu da kolleksiyada 2

fərqli qrupun mövcudluğunu göstərir. I qrup introduksiya edilmiş genotiplərin 68.5%-ini əhatə etmiş, yerli genotiplərin isə 82% -i II qrupda toplanmışdır. I və II qruplardakı genotiplərin orta əcdad qatqısı müvafiq olaraq 86.4% və 83.6% təşkil etmişdir. Klaster və PCoA analizlərinin nəticələri, STRUCTURE analizi ilə uyğunluq təşkil etməklə yerli və introduksiya olunmuş genotiplər arasındakı kəskin bir fərqliliyin mövcudluğunu göstərmişdir. Nümunələrin qruplaşma xarakterinə təsir edən digər amillər növmüxtəlifliyi əlamətləri və geneologiya olmuşdur. Tədqiqat çərçivəsində aşkar olunmuş TNP markerləri genotipləşdirmə və marker-əlamət asosiasiya analizləri üçün genetik mənbə olaraq istifadə ediləcəkdir. Əldə olunmuş verilənlər gələcək genetik analizlərdə və seleksiya üçün yeni strategiyaların hazırlanmasında və həyata keçirilməsində uğurla tətbiq edilə bilər.

Açar sözlər: *Yumşaq buğda, genom, sekvens əsasında genotipləşdirmə, TNP, tranzisiya, transversiya.*

Изучение геномной вариации в коллекции мягкой пшеницы на основе данных секвенирования нового поколения

З. Акпаров, М.А. Аббасов

Институт генетических ресурсов НАН Азербайджана

Генотипирование путем секвенирования (GBS) - это метод секвенирования следующего поколения (NGS), который использует рестрикционные ферменты для снижения сложности генома и успешно применяется при селекции растений. В настоящем исследовании было оценено геномное разнообразие 87 местных и интродуциованных генотипов мягкой пшеницы с использованием технологии GBS. Всего для трех геномов были получены 411 однонуклеотидных полиморфизмов (ОНП). Диапазон ОНП в геноме A, B и D варьировал в пределах 15–29, 10–36 и 3–17 соответственно. Наибольшее количество маркеров ОНП было установлено для генома B (48,8%), наименьшее – для генома D (14%). В целом, 70.2% ОНП оказались транзиционного (Ts), а 29.8% трансверсионного (Tv) типа. Наибольшее значение Delta K было зарегистрировано при K = 2, что указывает на наличие 2 различных групп в коллекции. I группа охватывала 68.5% интродуцированных образцов, тогда как 82% местных генотипов приходилось на II группу. Средний наследственный вклад генотипов в I и II группах составил 86.4% и 83.6% соответственно. Результаты кластерного и PCoA-анализов в соответствии со структурным анализом показали наличие резкой дифференциации между локальной и интродуцированной гермоплазмой. Другими факторами, определяющими характер группировки образцов, были признаки разновидностей и генеалогия. Маркеры ОНП, выявленные в настоящем исследовании, будут использоваться в качестве генетического источника для генотипирования и анализа ассоциаций маркер-признак. Полученные данные могут быть успешно применены при разработке и внедрении новых стратегий для последующих генетических анализов и селекции.

Ключевые слова: *Мягкая пшеница, геном, генотипирование путем секвенирования, ОНП, транзиция, трансверсия.*